

Application of Information Theory, Lecture 10

Hardcore Predicates

Handout Mode

Iftach Haitner

Tel Aviv University.

December 22, 2015

Part I

Motivation and Definition

Hardcore predicates

- ▶ Let $f: \{0, 1\}^n \mapsto \{0, 1\}^n$ be a “hard to invert” function, how unpredictable is x given $f(x)$
- ▶ Parts of x might be (totally) predictable
- ▶ It turns out that there is an hardcore part in x .

Hardcore predicates, cont.

Definition 1 (hardcore predicates)

A predicate $b: \{0, 1\}^n \mapsto \{0, 1\}$ is (s, ε) -hardcore predicate of $f: \{0, 1\}^n \mapsto \{0, 1\}^n$, if $\Pr_{x \leftarrow \{0, 1\}^n} [P(f(x)) = b(x)] \leq \frac{1}{2} + \varepsilon$, for any s -size P .

- ▶ Why size?
- ▶ We will typically consider poly-time computable f and b .
- ▶ Does every function has such a predicate?
- ▶ Does every hard to invert function has such a predicate?
- ▶ Is there a generic hardcore predicate for all hard to invert functions?

Let f be a function and let b be a predicate, then b is typically not a hard-core predicate of $g(x) = (f(x), b(x))$.

Part II

The Information Theoretic Settings

Some definitions

Let $f: \mathcal{D} \mapsto \mathcal{R}$.

- ▶ $\text{Im}(f) = \{f(x): x \in \mathcal{D}\}$.
- ▶ $f^{-1}(y) = \{x \in \mathcal{D}: f(x) = y\}$
- ▶ f is d regular, if $|f^{-1}(y)| = d$ for every $y \in \text{Im}(f)$.
- ▶ min entropy of $X \sim p$ is
$$H_{\infty}(X) = \min_{x \in \mathcal{X}} \{-\log p(x)\} = -\log \max_{x \in \mathcal{X}} \{p(x)\}.$$
- ▶ Examples:
 - ▶ Z is uniform over 2^k -size set.
 - ▶ $Z = X \mid_{f(X)=y}$, for 2^k -regular f , $y \in \text{Im}(f)$ and $X \leftarrow \mathcal{D}$.
- ▶ In both examples $H_{\infty}(Z) = k$

2-universal families

Definition 2 (2-universal families)

A function family $\mathcal{G} = \{g: \mathcal{D} \mapsto \mathcal{R}\}$ is **2-universal**, if $\forall x \neq x' \in \mathcal{D}$ it holds that $\Pr_{g \leftarrow \mathcal{G}} [g(x) = g(x')] = \frac{1}{|\mathcal{R}|}$.

Example: $\mathcal{D} = \{0, 1\}^n$, $\mathcal{R} = \{0, 1\}^m$ and $\mathcal{G} = \{A \in \{0, 1\}^{m \times n}\}$ with $A(x) = A \times x \bmod 2$.

Lemma 3 (leftover hash lemma)

Let X be a rv over $\{0, 1\}^n$ with $H_2(X) \geq k$ let $\mathcal{G} = \{g: \{0, 1\}^n \mapsto \{0, 1\}^m\}$ be 2-universal and let $G \leftarrow \mathcal{G}$. Then $SD((G, G(X)), (G, \sim \{0, 1\}^m)) \leq \frac{1}{2} \cdot 2^{(m-k)/2}$.

Hardcore predicate for regular functions

Lemma 4

Let $f: \{0, 1\}^n \mapsto \{0, 1\}^n$ be 2^k -regular function, let $\mathcal{G} = \{g: \{0, 1\}^n \mapsto \{0, 1\}\}$ be 2-universal and let $v: \{0, 1\}^n \times \mathcal{G} \mapsto \{0, 1\}^n \times \mathcal{G}$ be defined by $v(x, g) = (f(x), g)$.
Then $b(x, g) = g(x)$ is $(\infty, 2^{-(k-1)/2})$ hardcore-predicted of v .

- b is an hardcore predicate of v (not of f)

Proving Lemma 4

Claim 5

$SD((f(X), G, G(X)), (f(X), G, U)) \leq 2^{-(k-1)/2}$,
for $G \leftarrow \mathcal{G}$, $X \leftarrow \{0, 1\}^n$ and $U \leftarrow \{0, 1\}$.

We conclude the proof showing that indistinguishability implies unpredictability.

Lemma 6 (predicting to distinguishing)

Let (Y, Z) be rv over $\{0, 1\}^* \times \{0, 1\}$ and let P be an algorithm with $\Pr[P(Y) = Z] \geq \frac{1}{2} + \varepsilon$. Then \exists algorithm D , with essentially the same complexity as P , with $\Pr[D(Y, Z) = 1] - \Pr[D(Y, U) = 1] \geq \varepsilon$.

Proof: $D(y, z)$ outputs 1 if $P(y) = z$ and 0 otherwise. \square

Corollary 7

If $SD((Y, Z), (Y, U)) < \varepsilon$, then $\Pr[P(Y) = Z] < \frac{1}{2} + \varepsilon$ for **any** predictor P .

Proving Claim 5

For $y \in \text{Im}(f)$, let X_y be uniformly distributed over $f^{-1}(y)$.

Compute

$$\begin{aligned} & \text{SD}((f(X), G, G(X)), (f(X), G, U)) \\ &= \sum_{y \in \text{Im}(f)} \Pr[f(X) = y] \cdot \text{SD}((y, G, G(X))|_{f(X)=y}, (y, G, U)) \quad (\text{board}) \\ &= \sum_{y \in \text{Im}(f)} \Pr[f(X) = y] \cdot \text{SD}((y, G, G(X_y)), (y, G, U)) \\ &\leq \max_{y \in \text{Im}(f)} \text{SD}((y, G, G(X_y)), (y, G, U)) \\ &= \max_{y \in \text{Im}(f)} \text{SD}((G, G(X_y)), (G, U)) \end{aligned}$$

Since $H_\infty(X_y) = k$ for every $y \in \text{Im}(f)$, the leftover hash lemma yields that

$$\begin{aligned} \text{SD}((G, G(X_y)), (G, U)) &\leq \frac{1}{2} \cdot 2^{(1-H_\infty(X_y))} \\ &= 2^{(-k-1)/2}. \square \end{aligned}$$

Part III

The Computational Settings

Hard functions

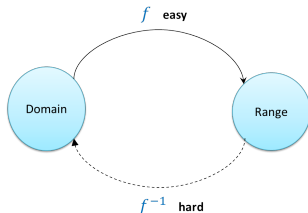
An injective function has hardcore bit, only if it is “hard to invert”.

Definition 8 (hard function)

$f: \{0, 1\}^n \mapsto \{0, 1\}^n$ is (s, ε) -hard, if
 $\Pr_{x \leftarrow \{0, 1\}^n} [\text{Inv}(f(x)) \in f^{-1}(f(x))] \leq \varepsilon$ for any s -size Inv .

- ▶ Size? Length preserving?
- ▶ f is hard \implies predicting x from $f(x)$ is hard.
- ▶ But does any hard function has an hardcore predicate?
- ▶ f is injective and not hard $\implies f$ has no hardcore predicate.

One-way functions



A one-way function (OWF) is:

- ▶ Easy to compute, **everywhere**
- ▶ Hard to invert, **on the average**
- ▶ Why should we care about OWFs?
- ▶ Hidden in (almost) **any** cryptographic primitive: necessary for "cryptography"
- ▶ Sufficient for many cryptographic primitives

One-way functions, cont.

Definition 9 (one-way functions (OWFs))

A poly-time $f: \{0, 1\}^n \mapsto \{0, 1\}^n$ is (s, ε) -one-way, if $\Pr_{x \leftarrow \{0, 1\}^n} [\text{Inv}(f(x)) \in f^{-1}(f(x))] \leq \varepsilon(n)$ for any $s(n)$ -size Inv .

- ▶ Poly-time computable: \exists poly-time algorithm such that $A(x) = f(x)$ for any $n \in \mathbb{N}$ and $x \in \{0, 1\}^n$.
- ▶ When restricted to input of length n , the function f is $(s(n), \varepsilon(n))$ -hard
- ▶ We typically consider $s = n^{\omega(1)}$ and $\varepsilon = n^{-\omega(1)}$.

The Goldreich-Levin predicate

For $x, r \in \{0, 1\}^n$, let $\langle x, r \rangle_2 := (\sum_{i=1}^n x_i \cdot r_i) \bmod 2 = \bigoplus_{i=1}^n x_i \cdot r_i$.

Theorem 10 (Goldreich-Levin)

For $f: \{0, 1\}^n \mapsto \{0, 1\}^n$, define $g: \{0, 1\}^n \times \{0, 1\}^n \mapsto \{0, 1\}^n \times \{0, 1\}^n$ by $g(x, r) = (f(x), r)$. Assume f is (s, ε) -hard, then $b(x, r) := \langle x, r \rangle_2$ is an $(\frac{\varepsilon}{n^2} \cdot s, \sqrt[3]{n\varepsilon})$ -hardcore predicate of g .

- ▶ Parameters are not tight, and we ignore small terms.
- ▶ If f is $(n^{\omega(1)}, 1/n^{\omega(1)})$ -hard, then b is an $(n^{\omega(1)}, 1/n^{\omega(1)})$ -hardcore predicate of g .
- ▶ Proof by reduction: a too small P for predicting $b(x, r)$ “too well” from $(f(x), r)$, implies a too small inverter for f :
- ▶ Assume $\exists s'$ -size P with $\Pr[P(g(X, R)) = b(X, R)] \geq \frac{1}{2} + \delta$, where hereafter R and X are iid uniformly distributed over $\{0, 1\}^n$
- ▶ We prove $\exists (\frac{n^2}{\delta^2} \cdot s')$ -size Inv with $\Pr[\text{Inv}(f(X)) = X] \in \Omega(\delta^3/n)$.

Focusing on a good set

Claim 11

There exists set $\mathcal{S} \subseteq \{0, 1\}^n$ with

1. $\frac{|\mathcal{S}|}{2^n} \geq \frac{\delta}{2}$, and
2. $\Pr[P(f(x), R) = b(x, R)] \geq \frac{1}{2} + \frac{\delta}{2}, \quad \forall x \in \mathcal{S}.$

Proof: Let $\mathcal{S} := \{x \in \{0, 1\}^n : \Pr[P(f(x), R) = b(x, R)] \geq \frac{1}{2} + \frac{\delta}{2}\}.$

$$\begin{aligned}\Pr[P(g(X, R)) = b(X, R)] &\leq \Pr[X \notin \mathcal{S}] \cdot \left(\frac{1}{2} + \frac{\delta}{2}\right) + \Pr[X \in \mathcal{S}] \\ &\leq \left(\frac{1}{2} + \frac{\delta}{2}\right) + \Pr[X \in \mathcal{S}]. \quad \square\end{aligned}$$

We conclude the theorem's proof showing that there exists a $\frac{n^2}{\delta^2}$ -size **Inv** with

$$\Pr[\text{Inv}(f(x)) = x] \in \Omega(\delta^2/n)$$

for every $x \in \mathcal{S}$. In the following we fix $x \in \mathcal{S}$.

The perfect case

$$\Pr [P(f(x), R) = b(x, R)] = 1$$



● $P(f(x), r) = b(x, r)$

● $P(f(x), r) \neq b(x, r)$

In particular, $P(f(x), e^i) = b(x, e^i)$ for every $i \in [n]$, for $e^i = (\underbrace{0, \dots, 0}_{i-1}, 1, \underbrace{0, \dots, 0}_{n-i})$.

Hence, $x_i = \langle x, e^i \rangle_2 = b(x, e^i) = P(f(x), e^i)$

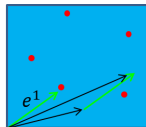
Algorithm 12 (Inverter Inv on input $y \in \text{Im}(f)$)

Return $(P(y, e^1), \dots, P(y, e^n))$.

$\text{Inv}(f(x)) = x$.

Easy case

$$\Pr[P(f(x), R) = b(x, R)] \geq 1 - \frac{1}{4n}$$



- $P(f(x), r) = b(x, r)$
- $P(f(x), r) \neq b(x, r)$

Fact 13

1. $b(x, w) \oplus b(x, y) = b(x, w \oplus y)$, for every $w, y \in \{0, 1\}^n$.
2. $\forall r \in \{0, 1\}^n$, the rv $(R \oplus r)$ is uniformly distributed over $\{0, 1\}^n$.

Hence, $\forall i \in [n]$:

1. $x_i = b(x, e^i) = b(x, r) \oplus b(x, r \oplus e^i)$ for every $r \in \{0, 1\}^n$
2. $\Pr[P(f(x), R) = b(x, R) \wedge P(f(x), R \oplus e^i) = b(x, R \oplus e^i)] \geq 1 - 2 \cdot \frac{1}{4n}$

Algorithm 14 (Inverter Inv on input y)

Return $(P(y, R) \oplus P(y, R \oplus e^1)), \dots, P(y, R) \oplus P(y, R \oplus e^n))$.

$$\Pr[\text{Inv}(f(x)) = x] \geq 1 - 2n \cdot \frac{1}{4n} = \frac{1}{2}$$

Proving Fact 13

1. For $w, y \in \{0, 1\}^n$:

$$\begin{aligned} b(x, y) \oplus b(x, w) &= \left(\bigoplus_{i=1}^n x_i \cdot y_i \right) \oplus \left(\bigoplus_{i=1}^n x_i \cdot w_i \right) \\ &= \bigoplus_{i=1}^n x_i \cdot (y_i \oplus w_i) \\ &= b(x, y \oplus w) \end{aligned}$$

2. For $r, y \in \{0, 1\}^n$:

$$\Pr[R \oplus r = y] = \Pr[R = y \oplus r] = 2^{-n}$$

Intermediate case

$$\Pr[P(f(x), R) = b(x, R)] \geq \frac{3}{4} + \frac{\delta}{2}$$



● $P(f(x), r) = b(x, r)$

● $P(f(x), r) \neq b(x, r)$

For any $i \in [n]$

$$\begin{aligned} & \Pr[P(f(x), R) \oplus P(f(x), R \oplus e^i) = x_i] \\ & \geq \Pr[P(f(x), R) = b(x, R) \wedge P(f(x), R \oplus e^i) = b(x, R \oplus e^i)] \\ & \geq 1 - \left(1 - \left(\frac{3}{4} + \frac{\delta}{2}\right)\right) - \left(1 - \left(\frac{3}{4} + \frac{\delta}{2}\right)\right) = \frac{1}{2} + \delta \end{aligned}$$

Algorithm 15 (Inv(y))

For every $i \in [n]$:

1. Sample $r^1, \dots, r^v \in \{0, 1\}^n$ uniformly at random
2. Let $m_i = \text{maj}_{j \in [v]} \{P(y, r^j) \oplus P(y, r^j \oplus e^i)\}$

Output (m_1, \dots, m_n)

Inv's success probability

The following claim holds for “large enough” v .

Claim 16

For every $i \in [n]$, it holds that $\Pr[m_i = x_i] \geq 1 - \frac{1}{2n}$.

Hence, $\Pr[\text{Inv}(f(x)) = x] \geq \frac{1}{2}$. Proof: (of claim):

- ▶ For $j \in [v]$, let W^j be 1, iff $P(f(x), r^j) \oplus P(f(x), r^j \oplus e^j) = x_j$.
- ▶ We need to lowerbound $\Pr\left[\sum_{j=1}^v W^j > \frac{v}{2}\right]$.
- ▶ W^j are iids and $E[W^j] \geq \frac{1}{2} + \delta$, for every $j \in [v]$

Lemma 17 (Hoeffding's inequality)

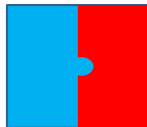
Let X^1, \dots, X^v be iids over $[0, 1]$ with expectation μ . Then,

$\Pr\left[\left|\frac{\sum_{j=1}^v X^j}{v} - \mu\right| \geq \alpha\right] \leq 2 \cdot \exp(-2\alpha^2 v)$ for every $\alpha > 0$.

- ▶ Hence, the proof follows for $v = \left\lceil \log(n) \cdot \frac{1}{2\delta^2} \right\rceil + 1$.

The actual (hard) case

$$\Pr[P(f(x), R) = b(x, R)] \geq \frac{1}{2} + \frac{\delta}{2}$$



● $P(f(x), r) = b(x, r)$

● $P(f(x), r) \neq b(x, r)$

- ▶ What goes wrong?
- ▶ $\Pr[P(f(x), R) \oplus P(f(x), R \oplus e^i) = x_i] \geq \delta$
- ▶ Hence, using a random guess does better than using P :-<
- ▶ Idea: guess the values of $\{b(x, r^1), \dots, b(x, r^v)\}$
(instead of calling $\{P(f(x), r^1), \dots, P(f(x), r^v)\}$)
- ▶ **Problem:** tiny success probability
- ▶ **Solution:** choose the samples in a **correlated** manner

Algorithm Inv

- ▶ For $\ell \in \mathbb{N}$ ($\approx \log \frac{n}{\delta}$, to be determined later), let $v = 2^\ell - 1$.
- ▶ In the following $\mathcal{L} \subseteq [\ell]$ stands for a **non empty** subset

Algorithm 18 (Inverter Inv on $y = f(x) \in \{0, 1\}^n$)

1. Sample uniformly (and independently) $t^1, \dots, t^\ell \in \{0, 1\}^n$
2. **Guess** the value of $\{b(x, t^i)\}_{i \in [\ell]}$
3. For all $\mathcal{L} \subseteq [\ell]$: set $r^\mathcal{L} = \bigoplus_{i \in \mathcal{L}} t^i$ and compute $b(x, r^\mathcal{L}) = \bigoplus_{i \in \mathcal{L}} b(x, t^i)$.
4. For all $i \in [n]$, let $m_i = \text{maj}_{\mathcal{L} \subseteq [\ell]} \{P(f(x), r^\mathcal{L} \oplus e^i) \oplus b(x, r^\mathcal{L})\}$
5. Output (m_1, \dots, m_n)

- ▶ Fix $i \in [n]$, and let $W^\mathcal{L}$ be 1 iff $P(f(x), r^\mathcal{L} \oplus e^i) \oplus b(x, r^\mathcal{L}) = x_i$.
- ▶ We need to lowerbound $\Pr \left[\sum_{\mathcal{L} \subseteq [\ell]} W^\mathcal{L} > \frac{v}{2} \right]$
- ▶ Problem: the $W^\mathcal{L}$'s are **dependent**!

Analyzing Inv's success probability

1. Let T^1, \dots, T^ℓ be iid and uniform over $\{0, 1\}^n$.
2. For $\mathcal{L} \subseteq [\ell]$, let $R^\mathcal{L} = \bigoplus_{i \in \mathcal{L}} T^i$.

Claim 19

1. $\forall \mathcal{L} \subseteq [\ell]$, $R^\mathcal{L}$ is uniformly distributed over $\{0, 1\}^n$.
2. $\forall w, w' \in \{0, 1\}^n$ and $\mathcal{L} \neq \mathcal{L}' \subseteq [\ell]$, it holds that $\Pr[R^\mathcal{L} = w \wedge R^{\mathcal{L}'} = w'] = \Pr[R^\mathcal{L} = w] \cdot \Pr[R^{\mathcal{L}'} = w'] = 2^{-2n}$.

Proof: (1) is clear. For (2), assume wlg. that $1 \in (\mathcal{L}' \setminus \mathcal{L})$.

$$\Pr[R^\mathcal{L} = w \wedge R^{\mathcal{L}'} = w']$$

$$= \sum_{(t^2, \dots, t^\ell) \in \{0, 1\}^{(\ell-1)n}} \Pr[(T^2, \dots, T^\ell) = (t^2, \dots, t^\ell)] \cdot \Pr[R^\mathcal{L} = w \wedge R^{\mathcal{L}'} = w' \mid (T^2, \dots, T^\ell) = (t^2, \dots, t^\ell)]$$

$$= \sum_{(t^2, \dots, t^\ell): (\bigoplus_{i \in \mathcal{L}} t^i) = w} \Pr[(T^2, \dots, T^\ell) = (t^2, \dots, t^\ell)] \cdot \Pr[R^{\mathcal{L}'} = w' \mid (T^2, \dots, T^\ell) = (t^2, \dots, t^\ell)]$$

$$= \sum_{(t^2, \dots, t^\ell): (\bigoplus_{i \in \mathcal{L}} t^i) = w} \Pr[(T^2, \dots, T^\ell) = (t^2, \dots, t^\ell)] \cdot 2^{-n}$$

$$= 2^{-n} \cdot 2^{-n} = \Pr[R^\mathcal{L} = w] \cdot \Pr[R^{\mathcal{L}'} = w']. \square$$

Pairwise independence variables

Definition 20 (pairwise independent random variables)

A sequence of rv's X^1, \dots, X^v is **pairwise independent**, if $\forall i \neq j \in [v]$ and $\forall a, b$, it holds that $\Pr[X^i = a \wedge X^j = b] = \Pr[X^i = a] \cdot \Pr[X^j = b]$.

- ▶ By **Claim 19**, $r^{\mathcal{L}}$ and $r^{\mathcal{L}'}$ (chosen by **Inv**) are pairwise independent for every $\mathcal{L} \neq \mathcal{L}' \subseteq [\ell]$.
- ▶ Hence, also $W^{\mathcal{L}}$ and $W^{\mathcal{L}'}$ are.
(Recall, $W^{\mathcal{L}}$ is 1 iff $P(f(x), r^{\mathcal{L}} \oplus e^i) \oplus b(x, r^{\mathcal{L}}) = x_i$)

Lemma 21 (Chebyshev's inequality)

Let X^1, \dots, X^v be pairwise-independent random variables with expectation μ and variance σ^2 . Then, for every $\alpha > 0$: $\Pr \left[\left| \frac{\sum_{j=1}^v X^j}{v} - \mu \right| \geq \alpha \right] \leq \frac{\sigma^2}{\alpha^2 v}$.

Inv's success provability, cont.

- Assuming that **Inv** always guesses $\{b(x, t^i)\}$ correctly, then $\forall \mathcal{L} \subseteq [\ell]$:

- $E[W^{\mathcal{L}}] \geq \frac{1}{2} + \frac{\delta}{2}$
 - $V(W^{\mathcal{L}}) := E[(W^{\mathcal{L}})^2] - E[W^{\mathcal{L}}]^2 \leq 1$

- Taking $v = 2n/\delta^2$ (hence $\ell = \lceil \log \frac{2n}{\delta^2} \rceil$), by Chebyshev's inequality for $i \in [n]$ it holds that

$$\Pr[m_i = x_i] = \Pr \left[\frac{\sum_{\mathcal{L} \subseteq [\ell]} W^{\mathcal{L}}}{v} > \frac{1}{2} \right] \geq 1 - \frac{1}{2n}.$$

- By a union bound, **Inv** outputs x with probability $\frac{1}{2}$.
- Taking the guessing probability into account, yields that **Inv** outputs x with probability at least $2^{-\ell}/2 \in \Theta(\delta^2/n)$.
- Recalling that we guaranteed to work well on $\frac{\delta}{2}$ of the x 's. We conclude that $\Pr[\text{Inv}(f(x)) = x] \in \Theta(\delta^3/n)$.

Reflections

- ▶ Hardcore functions:

Similar ideas allows to output $\log n$ "pseudorandom bits"

- ▶ Alternative proof for the leftover hash lemma:

Let X be a rv with over $\{0, 1\}^n$ with $H_\infty(X) \geq k$, and assume $SD((R, \langle R, X \rangle_2), (R, U)) > \alpha = 2^{-c \cdot k}$ for some universal $c > 0$.

$\implies \exists$ (a possibly inefficient) D that distinguishes $(R, \langle R, X \rangle_2)$ from (R, U) with advantage α

$\implies \exists P$ that predicts $\langle R, X \rangle_2$ given R with prob $\frac{1}{2} + \alpha$ (?)

\implies (by GL) $\exists \text{ Inv}$ that guesses X from nothing, with prob $\alpha^{O(1)} > 2^{-k}$

Reflections cont.

- List decoding:

- Encoder $f: \{0, 1\}^n \mapsto \{0, 1\}^m$ and decoder g , such that for any $x \in \{0, 1\}^n$ and c of hamming distance at most $(\frac{1}{2} - \delta)$ from $f(x)$: g examines $\text{poly}(1/\delta)$ symbols of c and outputs a $\text{poly}(1/\delta)$ -size list that whp contains x
- The code we used here is known as the **Hadamard** code

- LPN - learning parity with noise:

Given polynomially many samples of the form $(R_i, \langle x, R_i \rangle_2 + \theta)$, for $R_i \leftarrow \{0, 1\}^n$ and boolean $\theta_i \sim (\frac{1}{2} - \delta, \frac{1}{2} - \delta)$, find x .

- The difference comparing to Goldreich-Levin — no control over the R 's.